STAT 315: Conditional Expectation

Luc Rey-Bellet

University of Massachusetts Amherst

luc@math.umass.edu

April 8, 2025

Conditional probability models

Often probability models are described by conditional pdf rather than by joint PDF's:

Example 1: In a manufacturing process items are defective with probability p but the probability of defect P is itself a random variable e.g. P is uniform on $[0, \frac{1}{2}]$. If Y is the number of defective items in a batch of n then the joint pdf of Y and

$$f(k, p) = f(k|p)f(p) = \binom{n}{k} p^k (1-p)^{n-k} \times 2$$

with $k = 0, 1, 2, \cdots, n$ and $p \in [0, \frac{1}{2}]$. So when p is fixed the number of defective is binomial.

Example 2: The number of typos in a page are described by a Poisson random variable with parameter λ. The rate of typos Λ is itself random with a density $2e^{-2\lambda}$ (i.e. Λ is exponential with parameter $\beta = 1/2$.) So

$$f(n, \lambda) = f(n|\lambda)f(\lambda) = \frac{\lambda^n}{n!} e^{-\lambda} \times 2e^{-2\lambda}$$

Conditional RV

Recall the conditional PDF

$$p(y_1|y_2) = \frac{p(y_1, y_2)}{p(y_2)} \quad \text{discrete}$$

$$f(y_1|y_2) = \frac{f(y_1, y_2)}{f(y_2)} \quad \text{continuous}$$

For fixed y_2 they define a pdf since

$$\sum_{y_1} p(y_1|y_2) = \frac{\sum_{y_1} p(y_1, y_2)}{p(y_2)} = \frac{p(y_2)}{p(y_2)} = 1$$

$$\int_{-\infty}^{\infty} f(y_1|y_2) dy_1 = \frac{\int_{\infty}^{\infty} f(y_1, y_2) dy_1}{p(y_2)} = \frac{f(y_2)}{f(y_2)} = 1$$

Conditional RV

We write $Y_1|Y_2 = 2$ for the random variable with pdf $p(y_1|y_2)$ or $f(y_1|y_2)$.

Conditional Expectation

Conditional Expectation

The conditional expectation of Y_1 given $Y_2 = y_2$ is defined by

$$E[Y_1|Y_2 = y_2] = \begin{cases} \sum_{y_1} y_1\, p(y_1|y_2) & \text{discrete} \\ \int_{-\infty}^{\infty} y_1\, f(y_1|y_2) dy_1 & \text{continuous} \end{cases}$$

or more generally

For a function $g(Y_1)$ the conditional expectation of $g(Y_1)$ given $Y_2 = y_2$ is defined by

$$E[g(Y_1)|Y_2 = y_2] = \begin{cases} \sum_{y_1} g(y_1) p(y_1|y_2) & \text{discrete} \\ \int_{-\infty}^{\infty} g(y_1) f(y_1|y_2) dy_1 & \text{continuous} \end{cases}$$

The conditional expectation formula

Conditional Expectation Formula

Consider the function $g(y_2) = E[Y_1|Y_2 = y_2]$ then $g(Y_2) = E[Y_1|Y_2]$ is a random variable and we have the formulas

$$E[Y_1] = E\left[E[Y_1|Y_2]\right]$$

Proof: $g(y_2) = E[Y_1|Y_2 = y_2] = \int_{-\infty}^{\infty} y_1 f(y_1|y_2) dy_1$

Then using $f(y_1, y_2) = f(y_1|y_2) f(y_2)$ and $f(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2) dy_2$

$$E\left[E[Y_1|Y_2]\right] = E[g(Y_2)] = \int_{-\infty}^{\infty} g(y_2) f(y_2) dy_2$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_1 f(y_1|y_2) f(y_2) dy_1 dy_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_1 f(y_1, y_2) dy_1 dy_2$$

$$= \int_{-\infty}^{\infty} y_1 f(y_1) dy_1 = E[Y_1]$$

Example: Random sum of random variables

Imagine an car insurance company.

- The size of each claim has a gamma distribution with parameter $\alpha = 5$ and $\beta^2 = 6$. The average claim is $\alpha\beta = 30$.
- The number of claims in any given week has a Poisson distribution with parameter $\lambda = 100$.

Let us find the expected amount of money the car insruance needs to pay out every week. If N is the number of claim and X_i is the size of the claim. Then the payout is

$$X_1 + X_2 + \cdots X_N \quad \text{random sum of RVs}$$

Condition on $N = n$

$$E[X_1 + X_2 + \cdots X_N | N = n] = E[X_1 + X_2 + \cdots X_n] = n\alpha\beta$$

$$E[X_1+X_2+\cdots X_N] = E[E[X_1+X_2+\cdots X_N|N]] = E[N]\alpha\beta = \lambda\alpha\beta = 30,000$$

More examples

- A company has three departments 1,2,3 containing respectively 1000, 500, 250, with average salaries of \$50,000, \$70,000 and \$200,000. What is the average salary in the company.
- You roll a fair dice. If you roll 1-3 you win that amount and if you roll 4-6, you roll again and win that amount. What is your expected gain.
- Suppose Y is uniform on $[0,1]$ and conditioned on $Y = y$ X is exponential with parameter $y+1$. Compute $E[X]$.

Conditional Variance

It is natural to define the variance of $Y_1|Y_2 = y_2$ by

$$V[Y_1|Y_2 = y_2] = E[Y_1^2|Y_2 = y_2] - E[Y_1|Y_2 = y_2]^2$$

Formula for the variance

$$V[Y_1] = E[V[Y_1|Y_2] + V[E[Y_1|Y_2]]$$

Proof:

$$\begin{aligned} V[Y_1] &= E[Y_1^2] - E[Y_1]^2 \\ &= E[E[Y_1^2|Y_2]] - E[E[Y_1|Y_2]]^2 \\ &= \underbrace{E[E[Y_1^2|Y_2]] - E[E[Y_1|Y_2]^2]}_{E[V[Y_1|Y_2]]} + \underbrace{E[E[Y_1|Y_2]^2] - E[E[Y_1|Y_2]]^2}_{V[E[Y_1|Y_2]]} \\ &= E[V[Y_1|Y_2]] + V[E[Y_1|Y_2]] \end{aligned}$$